MACRECUR - DIAGNOSTICS FOR USE WITH MULTIPLE REGRESSIVE RECURSIVE RESIDUALS

Douglas M. Hawkins, Dept. of Applied Statistics

School of Statistics, University of Minnesota

Technical Report #562

May, 1991

MacRecur - Diagnostics for Use with Multiple Regression Recursive Residuals

Douglas M Hawkins
School of Statistics
University of Minnesota
St Paul, MN 55108

Introduction - Multiple regression and case diagnostics

The MacRecur code is an implementation of the recursive diagnostics discussed in Hawkins (1991), which should be read by anyone wanting to use these diagnostics in a formal way. (Readers should be warned however that the MacRecur option of 'Forward' computation is, largely for historical reasons, called 'Backward' computation and vice versa in Hawkins 1991). The following comments show how MacRecur can be used for exploratory investigations of the compatibility of the different 'cases' in a multiple regression data set.

Multiple regression is a statistical technique that is applied (and often misapplied) very widely. The statistical model underlying multiple regression incorporates several components - notably

- (i) that the regression model is linear in each of the predictors;
- (ii) that the true residuals follow a normal distribution;
- (iii) that the true residuals have constant variance;
- (iv) that the true residuals are statistically independent.

Departures from these assumptions can wreak havoc on the fitting of a multiple regression, and on the interpretation of the results from the regression. Particularly damaging are departures from linearity, and the occurrence of outliers. While multiple regression has been widely used since the advent of computers made its fitting possible in the 1960's, widespread recognition of the damaging consequences of departures from model, and of the need for checking for such departures, is much more recent. As a result of this, today it is standard practice for a conscientious user to supplement the multiple regression fit with scatter plots of the dependent variable against each of the predictors, and to perform some checks of the residuals. Among the latter, common practice includes plotting the residuals versus the fitted values, and making normal probability plots of the residuals.

It is also common to supplement the visual intuitive interpretation of these plots with some formal numeric measures. In particular, the area of <u>case diagnostics</u> produces various numeric measures of model departure associated with each of the 'cases' used in the multiple regression. Two such numeric measures are fundamental:-

The LEVERAGE, or POTENTIAL measures the capacity of the case to draw the multiple regression toward itself. A case has a high leverage if its predictor values are very different than the predictor values of the other cases in the data set;

The RESIDUAL measures the deviation of the dependent variable measured for that case from the regression line fitted to all cases. We call the property of fitting or not fitting the regression OUTLYINGNESS. For ease of interpretation, MacRecur provides this outlyingness diagnostic in the form of a transformed quantity which follows a standard normal distribution.

These two properties of a case are quite distinct. A case may have high leverage or low, and be outlying or inlying. Cases with high leverage are very important. If they are 'good' (ie conform to the regression model), then they are of great value since they help fix the multiple regression much more firmly than could be done without them. If they are 'bad' however (in particular if they have erroneous values for their dependent variable), then they can do great damage to the fitted regression.

A case which is outlying does not conform to the regression that fits the remainder of the cases. There are many possible reasons for this. For example one of the values recorded for that case could be an error; the outlier could however equally well be indicative of a curvilinear relationship rather than a linear; or a symptom of departure from the assumption of constant variance or that of normal distribution for the true residuals. Whatever may be the cause, outliers require some careful attention which certainly includes, but is not limited to, checking them for recording errors and temporarily removing them from the data set to see how they affect the conclusions drawn from the overall analysis.

From these two basic measures of a case, many more derived measures have been defined. One particularly useful derived measure is the Cook's distance. This measures the INFLUENCE of a case. A case's influence is a measure of the amount by which the fitted regression changes if that case is removed from the data set. If a case has a high influence, then the regression fit is substantially different depending on whether that case is included or excluded from the data set. Such cases should also be temporarily removed from the data set and the regression refit, to see in what way the interpretation of the data set changes depending on whether the case is used or not used.

Influence is a derivative property; for a case to have high influence requires at least a moderate degree of both leverage and outlyingness.

Recursive fitting

While these standard diagnostics are very commonly effective in trapping problems with the model, there is not as wide recognition as there should be that they can be misleading when the data set contains more than one departure from model. A data set that contains two severe outliers for example may produce unexceptional case diagnostics for both. A data set that has a curvilinear response and an outlier may similarly pass the conventional screens. This is because the model departures can mask each other. One possible solution is that of recursive fitting. A conventional multiple regression analysis fits the regression model using all the cases, and then computes diagnostics for all the cases. Recursive fitting operates completely differently. In this approach, you start out using only the first few cases in the data set - just enough to fit a regression line. In a regression problem with n cases, p predictors and an intercept, this subset will consist of cases 1,2,...,p+1. This subset is used to fit a regression line and that line used to compute case disgnostics for case number p+2. Then the subset is augmented by case p+2, the regression refit, and the fit used to compute diagnostics for case number p+3. This case is then added to the subset, the regression refitted, and diagnostics for case p+4 computed. The procedure continues in this way until finally diagnostics for case number n are computed.

On the face of it, this procedure is very wasteful. It replaces the fitting of a single multiple regression with fitting n-p regressions. The attraction however is that provided the 'problem' cases do not appear too early in the sequence, the early regressions will not be affected by them, and so will provide reliable diagnostics for the cases up to and including the problem cases.

Thus this recursive fitting will often provide more reliable diagnosis of difficult data sets with more than one departure from model. Recursive analysis can be used to provide case diagnostics for leverage, influence and outlyingness that are analogous to those obtained from full-sample case diagnostics, but with the attraction that problem cases can cloud the interpretation only of cases following them in the series, not those preceding them.

<u>Case and cumulative diagnostics</u>

Apart from these diagnostics for individual cases, recursive analysis opens up another possibility - that of diagnostics comparing the regression of each subset with the regression using the full sample. This provides <u>cumulative diagnostics</u>. A cumulative diagnostic for leverage compares the spread of the predictors on the first k cases with the spread of the predictors in the full data set. A cumulative diagnostic for influence compares the vector of regression coefficients for the first k cases with that for the full regression. MacRecur provides both diagnostics. Two additional cumulative diagnostics are provided - cumulative sum (cusum) charts for location and scale. The location cusum is a running total of the case 'outlier Z', corrected for an 'allowed' offset from zero. The scale cusum is a running total of a quantity computed from the case 'outlier Z' which tests for constancy of variance.

MacRecur provides on request any of five diagnostic plots - of leverage (both case and cumulative), of influence (both case and cumulative), of outlyingness (case only) and the cusums for location and for scale (both of which are cumulative diagnostics).

<u>The ordering of cases for recursive analysis</u>

Recursive analysis consists of fitting a sequence of regressions to the data, adding one case at a time. A natural question is in what order the cases are added. There are several logical choices. The first and most obvious (and the default used by MacRecur for the first set of diagnostics) is the order in which the cases occur in the input file. This is called the 'time' order. With this ordering, the recursive analysis is powerful for detecting situations in which the regression that fits the later cases differs in some way from that fitting the earlier cases.

Another useful ordering is by the size of the Y value predicted by the regression. This is called the 'Fitted value' ordering. With it, recursive analysis is effective in detecting many types of curvilinearity in the response of the dependent variable to the set of predictors. Another set of possibilities is to order by the values of any of the predictors. This ordering allows recursive analysis to detect nonlinearities in the regression function with respect to that predictor variable, as well as interactions between that predictor and the others. The tutorial paper of Galpin and Hawkins (1984) gives several examples of the use of recursive residuals (though not the other diagnostics) for detecting this type of departure.

Apart from the issue of how to order the cases, there is another issue of whether they should be ordered by increasing value or decreasing value of the sorting variable. Sorting by increasing values ('forward') provides analysis using the earlier cases as a basis against which to compare the later; sorting by decreasing value ('backward') makes the later cases the basis for comparison. The two orderings provide different and generally complementary information on the compatibility of the cases in the data set.

MacRecur supports ordering of cases by any predictor, by the fitted value, and by time order. In addition, cases may be ordered either forward or backward. The different ordering options are selected from a menu.

Data requirements

MacRecur reads a file of input data. Each line of the file consists of a case number, followed by the values of the predictors, and finally the value of the dependent variable. The first line of the file must contain two values - for the number of cases in the data set and the number of predictors. The values on the file are read in free format, and are separated by one or more blanks.

MacRecur writes all the values of the diagnostics produced in any run to a log file. This log file provides a useful permanent record of the graphics seen in exploring a data set.

Running MacRecur

To carry out a recursive analysis of a data set, first create the data file and an empty log file. When MacRecur is launched, it opens the 'Options' window and queries for the names of these two files using the standard file dialog menu. It then reads the data set, computing the recursive diagnostics in the 'time forward' order. When the diagnostics have been computed, the screen displays a message 'Diagnostics computed, ready to plot'. You may now use the mouse to make various selections using the Ordering and the Plot menus.

The Plot menu

The Plot menu is used to request a display of one of the five diagnostic plots computed by MacRecur. These are the Leverage, Influence, Outlier, Location cusum, and Scale cusum plots. If you select any of these, the 'Diagnostics' window comes to the front, and displays the requested plot. In all plots, case diagnostics are shown by open circles, and cumulative diagnostics by solid lines.

The <u>ordering</u> of the points is important as it does not necessarily correspond with the order in the data set. Recursive analysis is carried out by successively adding cases, starting with the case in position p+1 and going to the end. The plots show the diagnostics for the cases in the order in which they are added to the model, from left to right. Thus the initial default ordering is 'time forward':- the leftmost point plotted is the first case in the data set for which a recursive diagnostic can be computed. If the 'time backward' ordering is used, the leftmost point plotted corresponds to the case p+1 from the <u>end</u> of the data. With other orderings - for example by fitted value or by one of the predictors - the ordering on the plot need not bear any relationship to the ordering in the original data set.

Pointing and clicking on the plot

The main use of recursive analysis is to detect that an individual case or block of cases does not conform to the regression model A typical symptom of such a case would be that it shows a case influence much higher than the case influence of the other cases. If this is the case, then one would almost inevitably want to know how the picture changes if that case is removed from the data set. This is done by using the mouse to point to the plot symbol for that case and clicking. The program then reports two regression back to you. One is the 'first part' regression. This is the regression fitted to all cases PRECEDING the clicked-on case in the current ordering. The other is the current 'Full' regression - the multiple regression fitted to all cases other than those that have been temporarily deleted. In the reporting, the program gives the coefficients and standard errors associated with the predictors X(00), X(01), ... Of these, X(00) refers to the intercept of the regression, while X(01), X(02) ... are the predictors in order of appearance in the data.

When the regression reporting is done, the clicked-on case is temporarily deleted from the data set. The temporary deletion is done by moving that case to the end of the data set and decrementing the size of the data set by 1 for plotting purposes. It is important to note that case diagnostics for these temporarily-deleted cases are computed and are recorded in the log file; they are not however computed for the deleted cases in their natural position in the data set, but for the deleted cases being added to the data set after all other non-deleted cases.

If the mouse was very close to any point when it was clicked, then no deletion occurs. The window reports the regressions but leaves the data intact. Thus if you would like to see a 'first part' regression but do not want to delete any cases, position the mouse so that it is closer to the case where you want the cut that it is to any other case, but still not within about a character width. The program will then report the regressions but leave the data intact and ready for you to perform further plots or other operations.

When a case is deleted, the diagnostics need to be recomputed. The Options window gives a message that the data are being resorted and diagnostics recomputed; when this is done, the screen shows a message that the program is ready to plot diagnostics again. Very likely, you will want to look at some or all of the plot screens again to see whether there are other problem cases. If at this point a further case arises that you would like to temporarily delete, then you can point and click on that case and it will be added to the temporarily deleted list.

It is generally not sensible statistically to try to delete more than one case at a time, and MacRecur makes no provision for deleting more than one case from a point and click.

<u>The Ordering menu</u>

As mentioned, there are many different potential orderings of the data for recursive analysis, and these orderings give different and complementary pictures of the data. The Ordering menu allows the user to change the ordering. The options are Predictor, Fitted, and Time. If you select Predictor (one of the predictor variables), you will be asked to type in the number of the predictor. So if for example you want to order by the fourth predictor, you would type 4 and hit the Return key. Ordering by Fitted or Time requires no further input from you.

Also on the Ordering menu are the options to order Forward or Backward.

The Ordering menu also has an option Restore. This is used if, having explored the deletion of one or more cases, you want to un-delete the temporarily deleted case(s) and look at the full data set again.

The final option on the Ordering menu is Exit - this is used to terminate a MacRecur run.

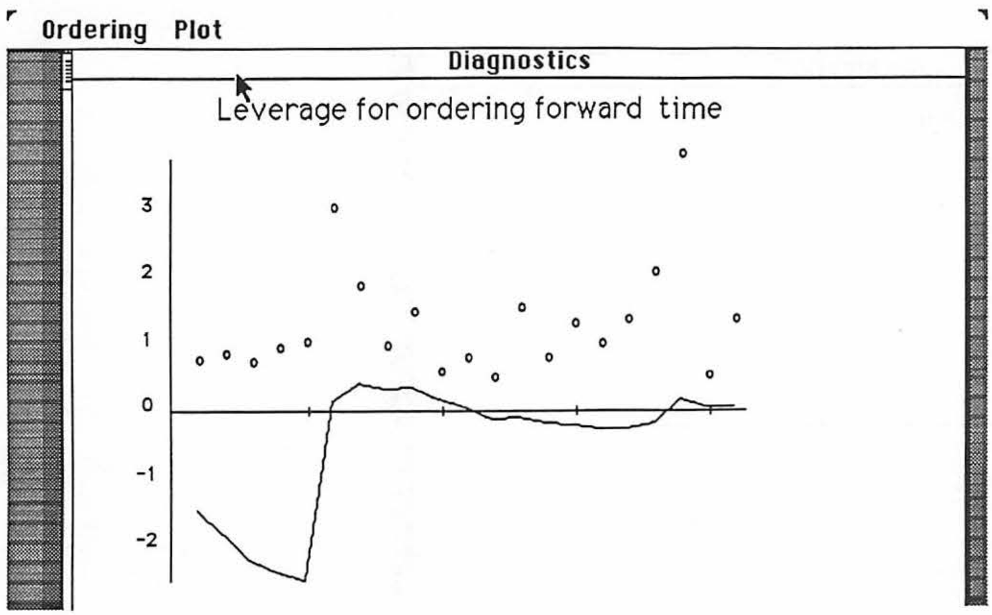
Example

To illustrate the use of MacRecur, we take the 'delivery' data set from Montgomery and Peck's 1982 book. This deals with the restocking of vending machines and records the dependent variable Y = the time taken to restock a machine from two predictors - the number of items stocked in the machine; and the distance traveled by the maintenance person to the machine. The data set is as follows:-

```
25   3
 1        7          560       16.68
 2        3          220       11.50
 3        3          340       12.03
 4        4           80       14.88
 5        6          150       13.75
 6        7          330       18.11
 7        2          110        8.00
 8        7          210       17.83
 9       30         1460       79.24
10        5          605       21.50
11       16          688       40.33
12       10          215       21.00
13        4          255       13.50
14        6          462       19.75
15        9          448       24.00
16       10          776       29.00
17        6          200       15.35
18        7          132       19.00
19        3           36        9.50
20       17          770       35.10
21       10          140       17.90
22       26          810       52.32
23        9          450       18.75
24        8          635       19.83
25        4          150       10.75
```

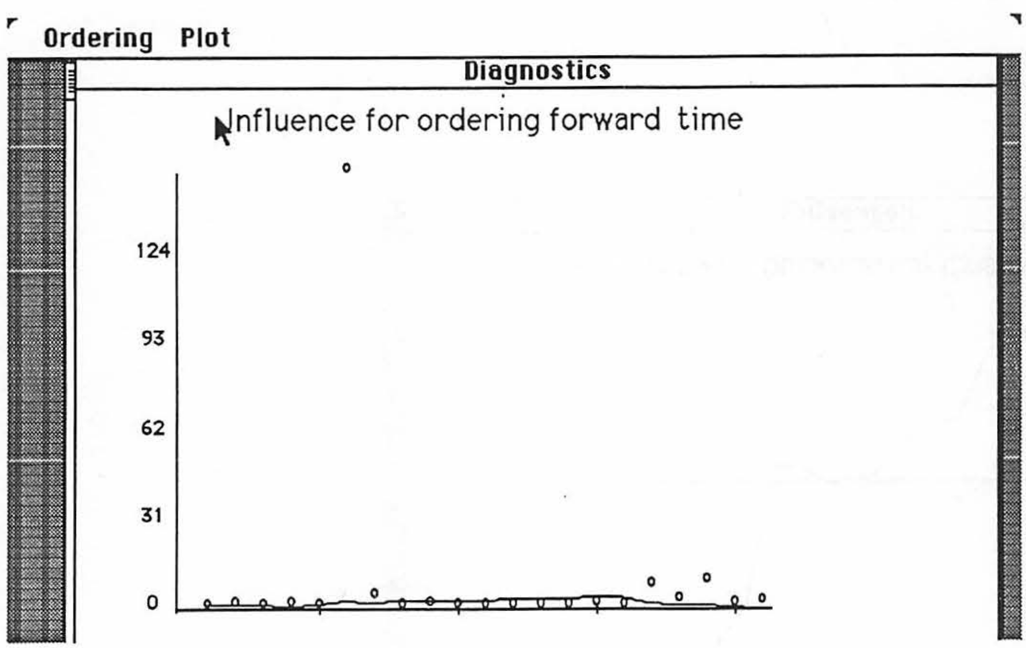
Conventional full-sample diagnostics show that case 9 is atypical. This is because it has by far the largest values for all three variables measured. Running MacRecur and producing all fives plots gives the following results:-

1. Leverage plot



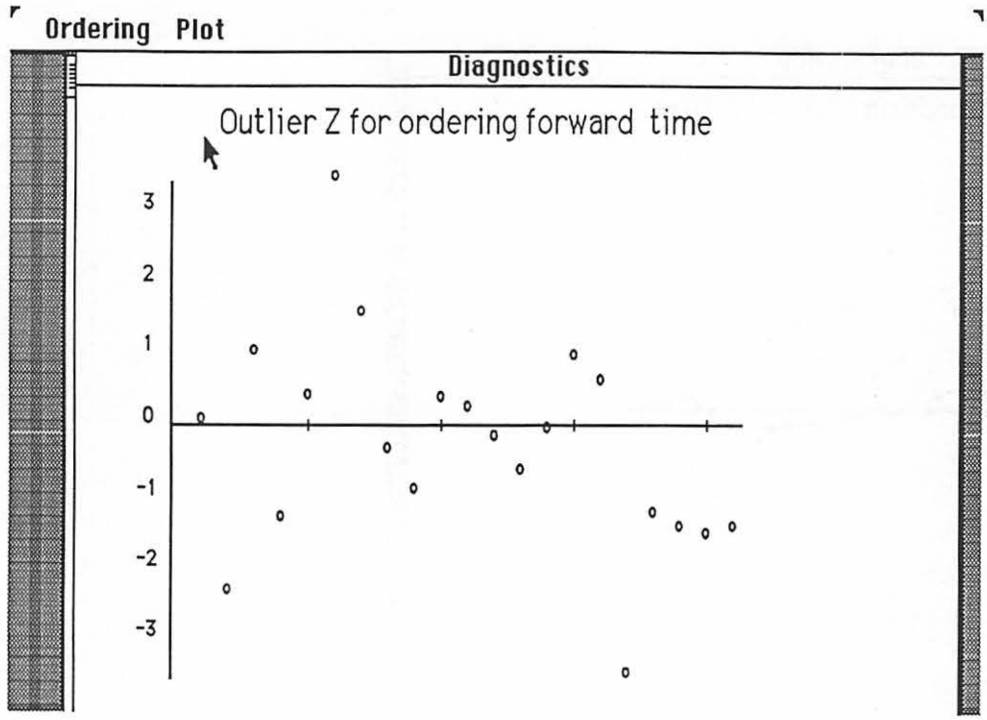
The sharp upward step of the cumulative leverage shows that the first cases have a far lower spread than does the entire sample. In addition, the sixth case and the third from last added have relatively high case leverages.

2. Influence plot



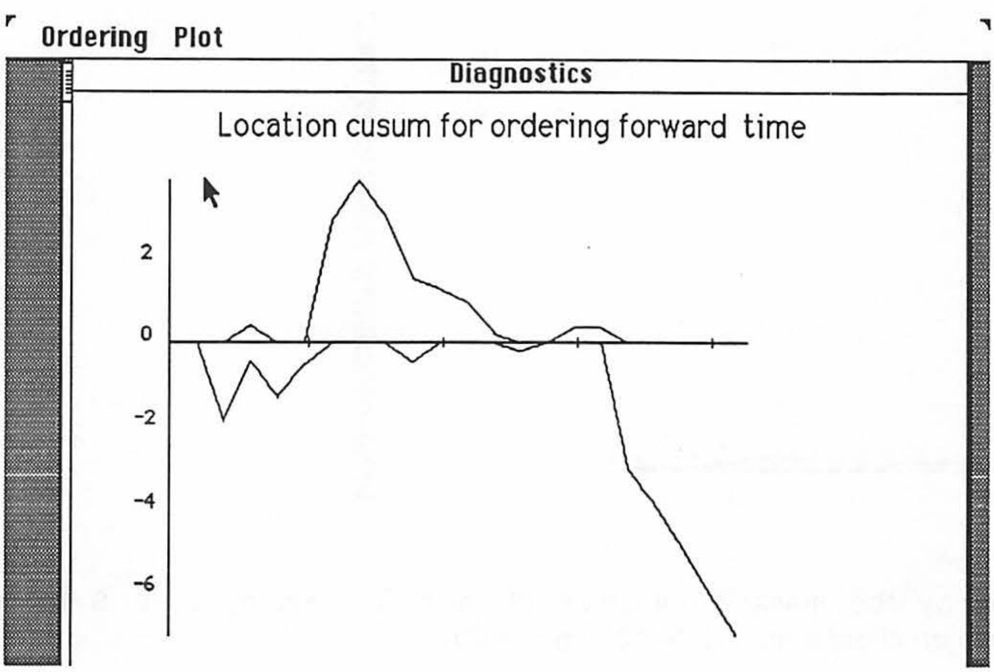
This plot is overwhelmed by the massive influence of case 9 - adding case 9 to its predecessors leads to a major change in the fitted regression.

3. Outlier plot



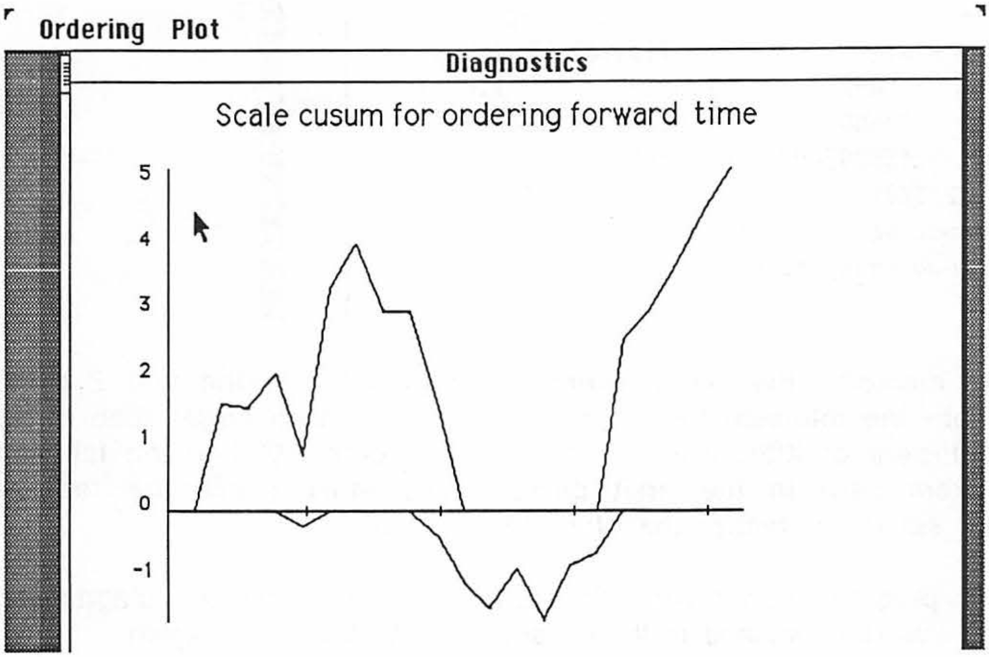
Each point tests whether the corresponding case is a significant outlier compared to the regression fitted to the preceding cases. The diagnostic follows a standard normal distribution if the data are good. The cases six from the left and five from the right appear to be outlying on the high and the low side respectively, though their Z values are not so large as to demand immediate attention.

4. Location cusum



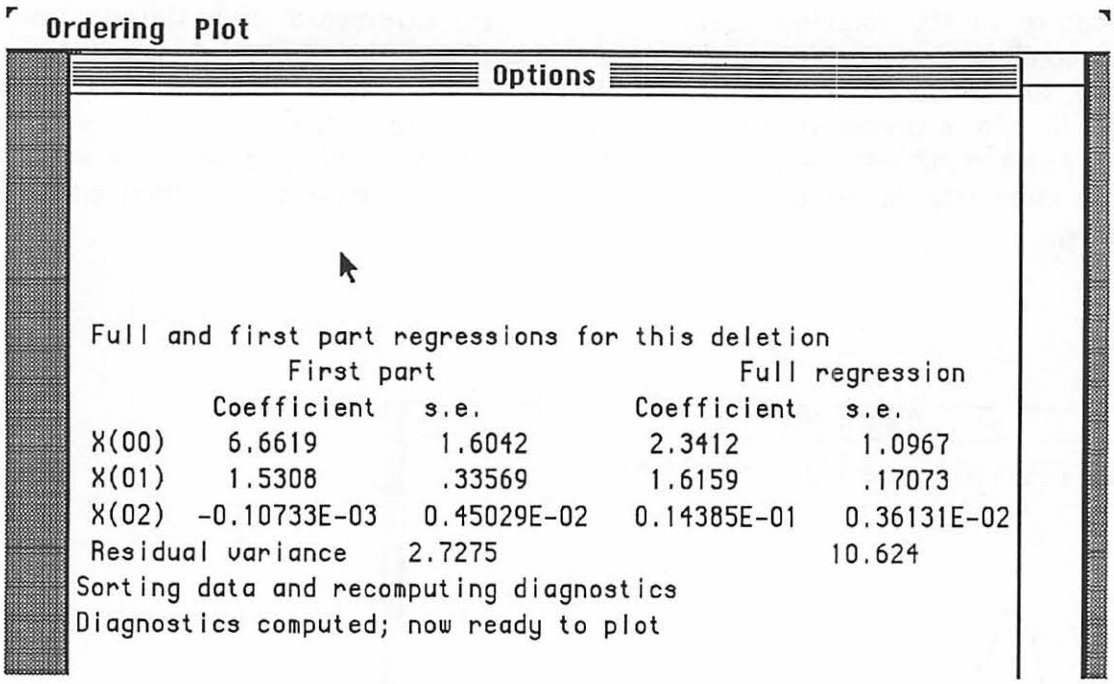
The noteworthy feature of the location cusum is the sharp downward drift starting about observation 21 and continuing to the end of the data. This shows that these last cases had a significant tendency to lie below the regression fitted to the earlier cases. This could indicate a change in the regression from the early to the last cases:- specifically the attendant servicing these machines may have been getting better with practice and so was able to perform the same job faster by the end of the data gathering period than was the case at the beginning.

5. Scale cusum



The rise in the scale cusum at the end of the plot is significant. It means that the variability of the last cases from the regression line exceeds that for the earlier cases.

These plots, taken together, indicate a lot of unanticipated structure in the data. The question might be - where to go next. This has a fairly easy answer - the case whose influence is so high as to make the rest if the influence plot structureless should be deleted to see what effect that has on the overall picture. Redraw the Influence plot, and point and click on this very influential case. The Diagnostics window goes to the back and the Options window to the front, and the program produces the following screen:-



Full and first part regressions for this deletion

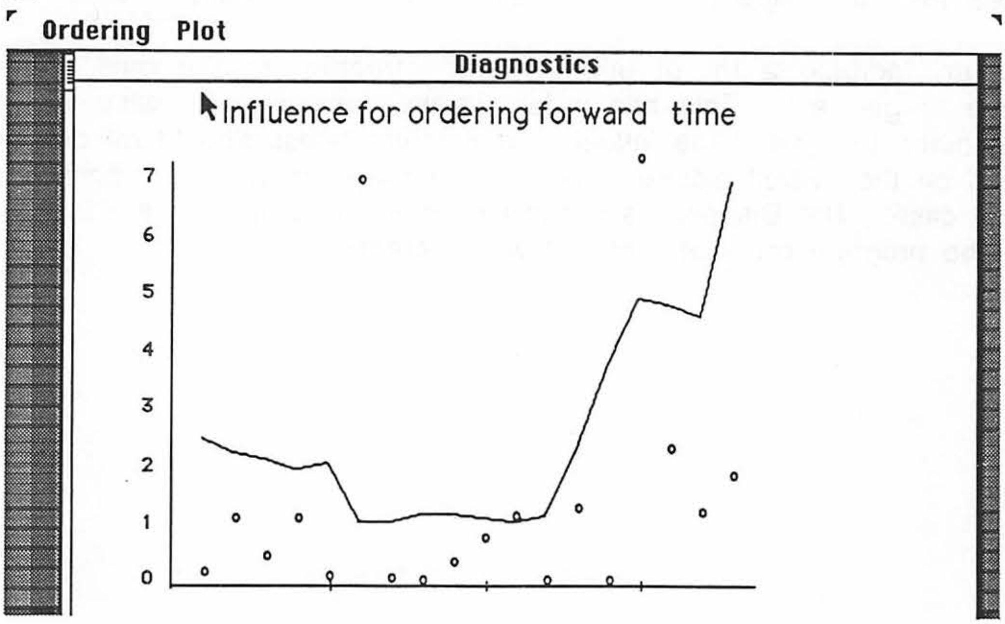
	First part		Full regression	
	Coefficient	s.e.	Coefficient	s.e.
X(00)	6.6619	1.6042	2.3412	1.0967
X(01)	1.5308	.33569	1.6159	.17073
X(02)	-0.10733E-03	0.45029E-02	0.14385E-01	0.36131E-02
Residual variance	2.7275		10.624	

Sorting data and recomputing diagnostics
Diagnostics computed; now ready to plot

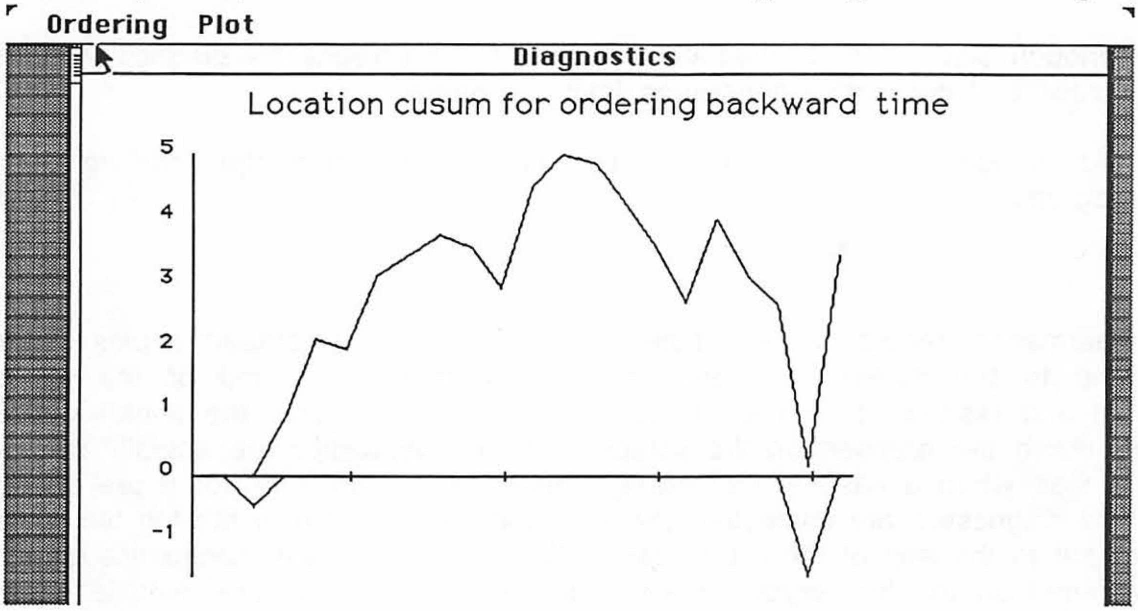
This certainly shows some marked difference between the regression to the first 8 cases and that to the full data set:- the intercept for the first 8 cases is much larger than in the full data set, while the coefficient of X(2), which is a very significant 0.014 in the full data set, is indistinguishable from zero in the 'frsit part'. Also striking are the residual variances - that for the full set is far higher than for the first part.

After deleting the case, the program then resorts the data and recomputes the diagnostics. It displays a message when this is done, and is then ready to plot diagnostics again.

When we replot the diagnostics, some more structure surfaces in the influence plot, which is:-



We can then do further exploration by testing the deletion of further points - for example these high influence cases which were previously overpowered by case 9. Alternatively however, we can try a different ordering. For example if we select the Backward option from the Ordering menu (ie getting the cases in reverse time order, with case 9 remaining deleted by transposition to the end of the data set), we get the following Location cusum:-



This shows that the recursive residuals have a mean that is systematically and substantially positive for about the first two thirds of the data, but that then drops and has values more regularly distributed around zero. We might be interested in seeing the regression as at the peak of the location cusum - this is the last of the sub regressions that appear to lie systematically below the data. Clicking on the peak produces the following Option screen:-

Ordering Plot

Options

Full and first part regressions for this deletion

	First part		Full regression	
	Coefficient	s.e.	Coefficient	s.e.
X(00)	3.9788	1.3413	4.4472	.95247
X(01)	1.5125	.16814	1.4977	.13021
X(02)	0.10878E-01	0.37481E-02	0.10324E-01	0.28536E-02
Residual variance	6.6763		5.9049	

Sorting data and recomputing diagnostics
Diagnostics computed; now ready to plot

(If we had wanted to see the two regressions but not delete any cases, then we would have clicked near but not too close to the peak. This would give us the two regressions but without the subsequent case deletion). The 'first part' does indeed seem to have an intercept systematically below that of the full regression - this is the apparent cause of the behavior of the location cusum.

This is probably enough discussion to illustrate the use of the screens for diagnosis of the model fit and the case and cumulative departures from the model.

Once you have had enough of exploring the data set, selecting Exit from the Ordering menu terminates the program.

The log file

The log file is a permanent record of the options tried in the run. It contains copies of the regressions echoed to the screen; a record of the orderings tried and of the cases temporarily deleted and restored to active status. Also valuable, it lists the actual values of the diagnostics which are graphed on the screen. These diagnostics are actually slightly more extensive in that, when a case is temporarily deleted, no diagnostics for it are shown on the screen. The diagnostics are computed however, and are recorded in the log file. The deleted cases are put to the end of the list of data. This means that the diagnostics quoted for them are computed on the full regression using all other cases. Where multiple cases have been deleted, they are added back when computing the diagnostics on a 'last out first in' basis. One might wonder why the diagnostics for these cases are computed and recorded but not plotted; the reason is scaling - that their diagnostics tend to be so large that they overshadow those of the other less atypical cases. The appendix records the log obtained in analyzing the delivery data.

Getting MacRecur

For those with access to electronic mail, an archive containing the executable program for MacRecur, this writeup, and the 'delivery' data set used as an example may be obtained from statlib. The archive is a StuffIt archive in BinHex format. To get it, send the one-line mail message (don't use a subject header)

send macrecur from general

to statlib@lib.stat.cmu.edu. You will require the UnStuffIt utility to extract the archive.

Acknowledgements

This work was performed using equipment and software supplied by Project MinneMac. The author was also supported by the National Science Foundation under grant DMS 9010983. Both sources of support are gratefully acknowledged.

References

Hawkins, D. M., (1991), 'Diagnostics for Use with Regression Recursive Residuals', Technometrics, 33, 221-234.

Galpin, J. S., and Hawkins. D. M., (1984), 'The use of recursive residuals in checking model fit in linear regression', Technometrics, 38, 94-105.

Montgomery, D. C., and Peck, E. A., (1982), 'Introduction to Linear Regression Analysis', Wiley, New York.

Appendix - log from run of delivery data

Case	Leverage case	Leverage cum	Influence case	Influence cum	Outlier Z	Loc cusum up	Loc cusum down	Scale cusum up	Scale cusum down
4	.685	-1.464	.141	1.753	.000	.00	.00	.00	-1.88
5	.778	-1.806	1.107	1.599	-2.412	.00	-1.91	1.62	.00
6	.662	-2.216	.424	1.519	.948	.45	-.46	1.56	.00
7	.873	-2.386	1.090	1.398	-1.370	.00	-1.33	2.07	.00
8	.964	-2.488	.104	1.479	.328	.00	-.51	.85	-.22
9	2.972	.152	154.392	3.185	3.428	2.93	.00	3.33	.00
10	1.786	.432	3.946	2.888	1.523	3.95	.00	4.03	.00
11	.887	.320	.129	3.663	-.418	3.03	.00	3.02	-.01
12	1.403	.367	1.167	3.415	-.980	1.55	-.48	3.00	.00
13	.521	.188	.030	3.412	.280	1.33	.00	1.65	-.35
14	.714	.062	.015	3.564	.162	.99	.00	.00	-1.06
15	.415	-.108	.021	3.725	-.261	.23	.00	.00	-1.47
16	1.462	-.079	.634	4.174	-.718	.00	-.22	.00	-.89
17	.725	-.178	.013	4.193	-.152	.00	.00	.00	-1.64
18	1.220	-.196	.739	4.455	.884	.38	.00	.00	-.80
19	.941	-.254	.188	4.595	.510	.39	.00	.00	-.61
20	1.289	-.263	7.437	2.516	-3.560	.00	-3.06	2.59	.00
21	2.003	-.175	2.837	2.127	-1.321	.00	-3.88	3.03	.00
22	3.782	.175	8.894	1.562	-1.509	.00	-4.89	3.71	.00
23	.458	.075	.767	.764	-1.601	.00	-5.99	4.50	.00
24	1.276	.065	2.128	.002	-1.507	.00	-7.00	5.17	.00
25	.736	.000	.003	.000	-.066	.00	-6.56	3.04	-1.14

Full and first part regressions for this deletion

	First part Coefficient	First part s.e.	Full regression Coefficient	Full regression s.e.
X(00)	6.6619	1.6042	2.3412	1.0967
X(01)	1.5308	.33569	1.6159	.17073
X(02)	-0.10733E-03	0.45029E-02	0.14385E-01	0.36131E-02
Residual variance	2.7275		10.624	

```
Case    9 (line    9) inactivated by interchange with case   25 (line  25)
   21 recursive residuals remain active.
```

Cases resorted forward by time

Case	Leverage case	Leverage cum	Influence case	Influence cum	Outlier Z	Loc cusum up	Loc cusum down	Scale cusum up	Scale cusum down
4	.685	-1.464	.141	2.521	.000	.00	.00	.00	-1.88
5	.778	-1.806	1.107	2.300	-2.412	.00	-1.91	1.62	.00
6	.662	-2.216	.424	2.185	.948	.45	-.46	1.56	.00
7	.873	-2.386	1.090	2.010	-1.370	.00	-1.33	2.07	.00
8	.964	-2.488	.104	2.128	.328	.00	-.51	.85	-.22
10	1.586	-2.245	6.983	1.132	3.074	2.57	.00	3.05	.00
11	2.895	-1.020	14.678	.791	1.409	3.48	.00	3.60	.00
12	1.337	-.981	.748	.584	-.782	2.20	-.28	3.28	.00
13	.516	-1.180	.008	.720	.141	1.84	.00	1.49	-.80
14	.709	-1.317	.070	.829	.356	1.70	.00	.34	-.95
15	.648	-1.456	.005	.939	.099	1.30	.00	.00	-1.92
16	2.002	-1.297	.011	.938	-.074	.72	.00	.00	-3.02
17	.685	-1.409	.033	1.083	-.254	.00	.00	.00	-3.44
18	1.140	-1.441	.814	2.351	.969	.47	.00	.00	-2.47
19	1.212	-1.460	.008	3.973	.090	.06	.00	.00	-3.48
20	3.133	-1.150	12.897	4.178	-2.195	.00	-1.70	1.41	-1.07
21	1.945	-1.064	3.038	3.955	-1.398	.00	-2.59	1.96	.00
22	5.096	-.398	.473	5.724	.242	.00	-1.85	.50	-.46
23	.496	-.499	1.050	5.661	-1.843	.00	-3.19	1.55	.00
24	1.396	-.496	2.075	5.126	-1.412	.00	-4.11	2.11	.00
25	.803	-.556	.167	5.050	-.520	.00	-4.13	1.32	.00
9	5.501	.000	79.266	.000	4.311	3.81	.00	4.46	.00

Cases resorted backward by time

Case	Leverage case	Leverage cum	Influence case	Influence cum	Outlier Z	Loc cusum up	Loc cusum down	Scale cusum up	Scale cusum down
22	.737	2.055	.789	2.955	.000	.00	.00	.00	-1.88
21	.747	1.643	.582	3.495	-1.003	.00	-.50	.02	-.86
20	.578	1.152	.128	6.384	-.514	.00	-.52	.00	-.67
19	.835	.953	1.185	6.190	1.566	1.07	.00	.74	.00
18	.623	.665	1.054	5.826	1.976	2.54	.00	1.94	.00
17	.478	.361	.053	5.799	.381	2.42	.00	.84	-.09
16	1.591	.539	4.530	5.175	1.991	3.91	.00	2.05	.00
15	.424	.292	.249	4.951	.942	4.36	.00	1.98	.00
14	.795	.164	.475	4.740	.899	4.75	.00	1.85	.00
13	.684	.021	.048	4.744	.302	4.56	.00	.56	-.29
12	.886	-.068	.007	5.089	-.102	3.95	.00	.00	-1.25
11	.945	-.140	2.696	4.012	2.286	5.74	.00	1.50	.00
10	1.747	-.053	1.843	4.425	1.120	6.36	.00	1.68	.00
8	.687	-.158	.063	4.505	.349	6.21	.00	.51	-.17
7	1.085	-.198	.036	4.531	-.203	5.51	.00	.00	-.75
6	.466	-.320	.005	4.660	-.123	4.88	.00	.00	-1.62
5	.810	-.393	.091	5.771	-.384	4.00	.00	.00	-1.70
4	.982	-.440	1.799	5.256	1.649	5.15	.00	.84	.00
3	1.109	-.471	.112	5.691	-.357	4.29	.00	.00	-.15
2	.806	-.534	0.000	5.690	.026	3.82	.00	.00	-1.56
1	1.168	-.556	2.785	4.877	-1.860	1.46	-1.36	1.07	.00
9	5.501	.000	79.266	.000	4.311	5.27	.00	4.21	.00

Full and first part regressions for this deletion

	First part Coefficient	First part s.e.	Full regression Coefficient	Full regression s.e.
X(00)	3.9788	1.3413	4.4472	.95247
X(01)	1.5125	.16814	1.4977	.13021
X(02)	0.10878E-01	0.37481E-02	0.10324E-01	0.28536E-02
Residual variance	6.6763		5.9049	

Case 10 (line 16) inactivated by interchange with case 1 (line 24)
20 recursive residuals remain active.

Cases resorted backward by time

Case	Leverage case	Leverage cum	Influence case	Influence cum	Outlier Z	Loc cusum up	Loc cusum down	Scale cusum up	Scale cusum down
4	.685	-1.464	.141	2.130	.000	.00	.00	.00	-1.88
5	.778	-1.806	1.107	1.943	-2.412	.00	-1.91	1.62	.00
6	.662	-2.216	.424	1.846	.948	.45	-.46	1.56	.00
7	.873	-2.386	1.090	1.699	-1.370	.00	-1.33	2.07	.00
8	.964	-2.488	.104	1.798	.328	.00	-.51	.85	-.22
11	2.507	-1.318	23.424	1.361	2.467	1.97	.00	2.52	.00
12	1.377	-1.241	2.471	.916	-1.518	.00	-1.02	3.20	.00
13	.543	-1.456	.063	.964	.392	.00	-.13	2.14	-.07
14	.995	-1.529	.583	.868	.863	.36	.00	1.95	.00
15	.648	-1.680	.036	.910	.270	.13	.00	.57	-.38
16	2.311	-1.407	.573	.910	.468	.10	.00	.00	-.28
17	.639	-1.537	.048	.988	-.317	.00	.00	.00	-.53
18	1.137	-1.570	.609	2.120	.828	.33	.00	.00	.00
19	1.137	-1.602	.002	3.316	.051	.00	.00	.00	-1.23
20	2.939	-1.303	13.034	3.666	-2.343	.00	-1.84	1.55	.00
21	2.047	-1.192	4.658	3.393	-1.715	.00	-3.06	2.47	.00
22	4.908	-.511	.026	3.799	.057	.00	-2.50	.28	-1.19
23	.530	-.613	1.027	3.864	-1.760	.00	-3.76	1.24	.00
24	1.642	-.578	1.729	4.863	-1.155	.00	-4.42	1.47	.00
25	.769	-.645	.151	4.805	-.509	.00	-4.43	.66	.00
10	2.155	-.556	4.343	4.267	1.585	1.09	-2.34	1.43	.00
9	5.501	.000	79.266	.000	4.311	4.90	.00	4.56	.00

All cases restored to active status

Cases resorted backward by time

Case	Leverage case	Leverage cum	Influence case	Influence cum	Outlier Z	Loc cusum up	Loc cusum down	Scale cusum up	Scale cusum down
3	.314	.987	.010	3.530	.000	.00	.00	.00	-1.88
4	1.010	1.446	.808	3.876	-.524	.00	-.02	.00	-1.67
5	.807	1.192	1.280	3.130	-1.961	.00	-1.49	1.18	.00
6	.353	.678	.057	3.043	-.478	.00	-1.46	.30	.00
7	.488	.324	.030	3.026	.282	.00	-.68	.00	-.34
8	.802	.152	.206	2.972	-.553	.00	-.73	.00	-.07
10	1.786	.432	3.946	2.695	1.523	1.02	.00	.69	.00
11	.887	.320	.129	3.418	-.418	.11	.00	.00	-.01
12	1.403	.367	1.167	3.186	-.980	.00	-.48	.00	.00
13	.521	.188	.030	3.184	.280	.00	.00	.00	-.35
14	.714	.062	.015	3.325	.162	.00	.00	.00	-1.06
15	.415	-.108	.021	3.476	-.261	.00	.00	.00	-1.47
16	1.462	-.079	.634	3.895	-.718	.00	-.22	.00	-.89
17	.725	-.178	.013	3.913	-.152	.00	.00	.00	-1.64
18	1.220	-.196	.739	4.157	.884	.38	.00	.00	-.80
19	.941	-.254	.188	4.287	.510	.39	.00	.00	-.61
20	1.289	-.263	7.437	2.348	-3.560	.00	-3.06	2.59	.00
21	2.003	-.175	2.837	1.985	-1.321	.00	-3.88	3.03	.00
22	3.782	.175	8.894	1.458	-1.509	.00	-4.89	3.71	.00
23	.458	.075	.767	.713	-1.601	.00	-5.99	4.50	.00
24	1.276	.065	2.128	.001	-1.507	.00	-7.00	5.17	.00
25	.736	.000	.003	.000	-.066	.00	-6.56	3.04	-1.14